Initial Value Problems in Domain Theory

A. Edalat
Department of Computing, Imperial College, UK

D. Pattinson
Institut für Informatik, LMU München, Germany

Abstract

We present a domain-theoretic version of the Picard operator and of Picard's theorem for solving classical initial value problems. Our formulation of the Picard operator allows us to compute solutions as least fixed points on the space of Scott continuous interval-valued maps of a real variable. In this setup, which considerably simplifies earlier approaches, we obtain fast convergence to the solution, given that the vector field is Lipschitz and can be approximated by step functions. Since the naive algorithm induced by the theorem suffers from a complexity explosion, the method is further refined, and we show that we can avoid the complexity explosion, while retaining the same speed of convergence. Finally, we report the performance of a simple implementation of our framework, tested on a few examples.

1 Introduction

We consider the initial value problem (IVP) of the form

$$f'(x) = v(x, f(x)), \quad f(0) = 0, \tag{1}$$

where $v : O \to \mathbb{R}^2$ is continuous in a neighbourhood $O \subseteq \mathbb{R}^2$ of the origin $(0,0) \in O$. By a theorem of Osgood there are always a lower and and an upper solutions [9]. Uniqueness of the solution is guaranteed, by Picard's theorem, if v satisfies a Lipschitz condition in its second argument. The question of computability and theoretical complexity of the initial value problem has been studied in different contexts in computable analysis [13, 3, 8, 15, 20, 17, 6].

On the algorithmic and more practical side, standard numerical packages for solving IVP try to compute an approximation to a solution with a specified degree of accuracy. Although these packages are usually robust, their methods are not guaranteed to be correct and it is easily to find examples where they output inaccurate results [14]. Interval analysis [18] provides a method to give upper and lower bounds for the unique solution in the Lipschitz case with a prescribed tolerance, and has been developed and implemented for analytic vector fields [19, 1].

Domain theory presents an alternative technique, which is based on proper data types, to produce a solution with a given degree of accuracy. Using an order-theoretic framework for differential calculus [12], in particular a domain for C^1 functions and a domain-theoretic Picard like theorem, a method for solving differential equations with a general initial condition was introduced in [11]. By expressing the vector field as supremum of step functions and successively updating the function and the derivative approximations, represented by consistent pairs of step functions in the domain of C^1 functions, approximations to solutions are found that are bounded below and above by, respectively, a lower and an upper semi-continuous functions, equivalently an interval-valued Scott continuous map. With a suitable selection of the initial Scott continuous map, it can also solve the classical IVP above. A linear algorithm for function updating and a quadratic algorithm for derivative updating of the consistent pairs of linear and simple step functions were presented in [11]. However, the method results in an exponential blow-up of the number of step functions used in the approximation.

In this paper, we derive a simplified domain-theoretic version of Picard's theorem for solving the classical IVP, which only uses the domain for C^0 functions, namely the function space of Scott continuous interval-valued maps of a real variable, with the pointwise partially ordering of maps. This new treatment is much more similar to the classical theorem and brings domain theory closer to the branch of analysis commonly called differential inclusions [5, 7], as the Scott continuous maps are precisely the upper semi-continuous maps in the well-established terminology of set-valued function theory. The major consequence of the simplified framework is that it gives rise, in the Lipschitz case, to fast convergence of the approximations to the solution, as is the case in the classical theorem. Moreover, we show that by flattening the linear step function into a simple step function, after each iterate of the Picard operator, and by a suitable approximation of the vector field, one can overcome the blow-up of the number of single-step functions. We illustrate with a few examples of IVP based on a simple implementation of the technique. As for future work, the framework will be extended to higher dimensions and a more refined implementation will be compared in scope of application and performance with existing interval analysis packages like AWA [1].

2 Preliminaries and Notation

For the remainder of the paper, we fix a continuous vector field

$$v : [-a, a] \times [-K, K] \to \mathbb{R}$$

which is defined in a rectangle containing the origin and consider the IVP given by 1. In order to guarantee that the expression $v(x, f(x))$ is well-defined for $x \in [-a, a]$ and any solution $f : [-a, a] \to \mathbb{R}$, we assume that $[-aM, aM] \subseteq (-K, K)$. Geometrically, M will be a bound of the derivative of a solution f. Since f also satisfies the initial condition $f(0) = 0$, we have that $|f(x)| \leq Mx$ for all $x \in [-a, a]$. We can therefore guarantee that the expression $v(x, f(x))$ is defined for all $x \in [-a, a]$ if $aM \leq K$. We require the stronger condition $[-aM, aM] \subseteq (-K, K)$ since we also need to accommodate approximations of the vector field v; this extra amount of space is denoted by $\delta = K - aM$.

In the sequel, we are going to consider interval-valued functions. These functions take values and some of their arguments in the interval domain

$$\mathbb{IR} = \{[a, b] \mid a, b \in \mathbb{R}, a \leq b\} \cup \{\mathbb{R}\}$$

where the order is given by reverse inclusion; the way-below relation $\ll$ is given by $[a, b] \ll [c, d]$ iff $[c, d] \subseteq (a, b)^o$. For a compact interval $[a, b]$, we denote the sub-domain of intervals contained in $[a, b]$ by $\mathrm{I}[a, b]$.

Every interval valued function $f : X \to \mathbb{IR}$ is given by an upper and a lower function; we write $f = [f^+, f^-]$ if $f(x) = [f^-(x), f^+(x)]$ for all $x \in X$ for this representation. Interval valued functions can be obtained by extending continuous functions $f : [a, b] \to \mathbb{R}$ to the domain of intervals, yielding

$$\hat{f} : \mathrm{I}[a, b] \to \mathbb{IR}, \quad A \mapsto [\inf_{x \in A} f(x), \sup_{x \in A} f(x)],$$

where A denotes a compact interval $A \subseteq [a, b]$; we often call this the canonical extension of f. It is easy to see that $\hat{f}$ is continuous wrt. the Scott topology on $\mathrm{I}[a, b]$ and $\mathbb{IR}$ if f itself is continuous. In particular, we can extend v canonically to an interval-valued function

$$\hat{v} : [-a, a] \times \mathrm{I}[-K, K] \to \mathbb{IR}, \quad (x, Y) \mapsto [\inf_{y \in Y} v(x, y), \sup_{y \in Y} v(x, y)].$$

Note that $[-a, a] \times \mathrm{I}[-K, K] \to \mathbb{IR}$, the collection of Scott continuous maps with the pointwise ordering, is a continuous Scott domain. We often identify a function (resp. vector field) with its canonical extension if it is clear from the context which we mean.

In the sequel, we will approximate both functions and vector fields by means of step functions. In order to measure the speed of convergence, as well as for technical convenience in the formulation of some of our results, we introduce the following notation:

1. If $f : [-a, a] \to \mathbb{IR}$ with $f(x) \neq \mathbb{R}$, we define the *width* of f by $w(f) = \sup\{f^+(x) - f^-(x) \mid x \in [-a, a]\}$.
2. For $u : [-a, a] \times \mathrm{I}[-K, K] \to \mathbb{IR}$ with $u \sqsubseteq v$, the *width of u relative to v* is defined as $w_v(u) = \sup_{Y \in \mathrm{I}[-K,K]} w(u(\cdot, Y)) - w(v(\cdot, Y))$.

Considering u as approximation to v, the relative width $w_v(u)$ can be understood as the quality of the approximation, or the distance between u and v.

3 The Picard Operator in Domain Theory

In the classical proof of Picard's theorem on the existence and uniqueness of the solution of the initial value problem (1) one defines an integral operator on $C^0[-a, a]$ by

$$f \mapsto \lambda t. \int_0^t v(x, f(x))dx$$

which can be shown to be a contraction for sufficiently small a provided v satisfies a Lipschitz condition in the second argument [16]. An application of Banach's fixed point theorem then yields a solution of the initial value problem. We now define the domain-theoretic Picard operator for arbitrary continuous functions $u : [-a, a] \times \mathrm{I}[-K, K] \to \mathbb{IR}$ and focus on the special case where u is the canonical extension of a classical function later.

Definition 3.1. Suppose $u : [-a, a] \times \mathrm{I}[-K, K] \to \mathbb{IR}$ is continuous. We define the Picard operator $P_u : ([-a, a] \to \mathbb{IR}) \to ([-a, a] \to \mathbb{IR})$ for $f = [f^-, f^+] : [-a, a] \to \mathbb{IR}$ by

$$P_u(f)(t) = [\int_0^t u^-(x, f(x))dx, \int_0^t u^+(x, f(x))dx]$$

in case both integrals are defined; otherwise $P_u(f)(t) = \mathbb{R}$.

Since u and f are Scott continuous, it follows that the functions $\lambda x.u^-(x, f(x))$ and $\lambda x.u^+(x, f(x))$ are, respectively lower and upper semi-continuous functions and thus measurable. Hence P_u is well defined.

Lemma 3.2. *If $u : [-a, a] \times \mathrm{I}[-K, K] \to \mathbb{IR}$ is Scott continuous, then so is P_u.*

Proof. Monotonicity of P_u is straightforward. If $(f_i)_{i\in\mathbb{N}}$ is a increasing sequence of functions $f_i : [-a, a] \to \mathbb{IR}$, then by continuity of u we have: $\lambda x.u(x, \bigsqcup_{i\in\mathbb{N}} f_i(x)) = \bigsqcup_{i\in\mathbb{N}} \lambda x.u(x, f_i(x))$ and the continuity of P_u follows from the monotone convergence theorem. $\square$

In the classical proof of Picard's theorem, one can choose an arbitrary function as initial approximation of the solution. In domain theory, we begin the approximation with the function that contains the least possible amount of information: in our case this is the function $\lambda t.[-K, K]$. This gives rise to a sub-domain of $[-a, a] \to \mathbb{IR}$, in which the solutions are approximated:

Notation 3.3. In the following, $f_0 : [-a, a] \to \mathbb{IR}$ denotes the function defined by $t \mapsto [-K, K]$. The associated upper set $\uparrow f_0 = \{f : [-a, a] \to \mathbb{IR} \mid f_0 \sqsubseteq f\}$ is denoted by D. Note that D is a sub-dcpo of $[-a, a] \to \mathbb{IR}$ with least element f_0. Finally, V denotes the set of Scott continuous functions $u : [-a, a] \times \mathrm{I}[-K, K] \to \mathbb{IR}$ with $w_v(u) \le \delta/a$.

The restriction on the relative width of $u \in V$ is needed to show that the Picard operator is well defined, which is the content of the following Lemma.

Lemma 3.4. *Suppose $u \in V$, $f \in D$.*

1. *$|u^\pm(x, Y)| \le K/a$ for all $(x, Y) \in [-a, a] \times \mathrm{I}[-K, K]$.*
2. *$(P_u(f))^\pm$ satisfies a Lipschitz condition with Lipschitz constant K/a.*
3. *$|(P_u(f))^\pm(x)| \le K$ for all $x \in [-a, a]$.*

Proof. Suppose $x \in [-a, a]$ and $Y \in \mathrm{I}[-K, K]$. Since $w_v(u) \le \delta/a$, we have $u^+(x, Y) - u^-(x, Y) - v^+(x, Y) - v^-(x, Y) \le \delta/a$, hence $|u^\pm(x, Y) - v^\pm(x, Y)| \le \delta/a$, which implies the claim, since $|v^\pm(x, Y)| \le M$ by construction and $K/a = M + \delta/a$.

For the Lipschitz condition, we have

$$\begin{aligned} |P_u(f)^\pm(t) - P_u(f)^\pm(t')| &\le \int_{t'}^{t} |u^\pm(x)| dx \\ &\le (M + \delta/a)|t - t'| = K/a|t - t'|. \end{aligned}$$

Finally, since $P_u(f)^{\pm}(0) = 0$, we obtain $|P_u(f)^{\pm}(t)| = |P_u(f)^{\pm}(t) - P_u(f)^{\pm}(0)| \leq K/a \cdot t \leq K$ for all $t \in [-a, a]$. ☐

As a corollary, we obtain:

Corollary 3.5. *Suppose $u \in V$. Then $P_u : D \to D$ is well defined and continuous.*

Proof. That P_u is well defined follows from $|P_u(f)^{\pm}(x)| \leq K$ for all $x \in [-a, a]$. Continuity has been established in Lemma 3.2.

☐

The last lemma puts us in the position to replace Banach's theorem by the Knaster-Tarski theorem in the process of constructing a solution of the initial value problem (1); recall that f_0 is the least element of D.

Theorem 3.6. *Suppose $f_{n+1} = P_v(f_n)$. Then $f = \bigsqcup_{n \in \mathbb{N}} f_n$ satisfies $P_v(f) = f$.*

Proof. Follows immediately from the Knaster Tarski Theorem, see e.g. [4, Theorem 2.1.19]. ☐

The bridge between the solution of the domain theoretic fixpoint equation and the classical initial value problem is established in the following lemma.

Lemma 3.7. *Suppose $f = [f^-, f^+] : [-a, a] \to \mathrm{I}[-K, K]$ satisfies $P_v(f) = f$ and $f^- = f^+$. Then $f^- = f^+$ solves (1).*

Proof. Since $f^- = f^+$, we can identify both functions with f. Since f is Scott continuous, we have that f is both upper and lower semi continuous, hence continuous. Furthermore, $f = P_v(f)$ implies that $f(t) = \int_0^t v(x, f(x))dx$, hence f is continuously differentiable and the claim follows from the fundamental theorem of calculus. ☐

In order to obtain a solution of the classical problem, we therefore need to find a fixpoint of P_v with width 0. This is the content of the following section, where we construct a zero width fixpoint by imposing a Lipschitz condition on v.

4 The Lipschitz Case

In order to obtain a solution to the classical problem, we impose the following Lipschitz condition on the defining vector field v:

Assumption 4.1. There is $L > 0$ and $0 < c < 1$ such that $aL < c < 1$ and $|v(x,y) - v(x,y')| \le L|y - y'|$ for all $(x,y) \in [-a,a] \times [-K,K]$.

The additional condition $aL < 1$ can always be ensured by restricting the domain of definition of v; this is as in the classical proof. Assuming the Lipschitz condition, we have the following estimate, which guarantees that the least fixed point of P_v is of width 0:

Lemma 4.2. *Suppose $f_0 \sqsubseteq f$. Then $w(P_v(f)) \le aL \cdot w(f)$.*

Proof. Using the Lipschitz condition, we calculate

$$\begin{aligned}
w(P_v(f)) &= \sup_{t\in[-a,a]} \int_0^t v^+(x,f(x)) - v^-(x,f(x))dx && \text{(Def'n of } P_v\text{)}\\
&= \sup_{t\in[-a,a]} \int_0^t \sup_{y\in f(x)} v^+(x,y) - \inf_{y\in f(x)} v^-(x,y)dx && \text{(Def'n of canonical extension)}\\
&\le \sup_{t\in[-a,a]} \int_0^t L|f^+(x) - f^-(x)|dx && \text{(Lipschitz condition on } v\text{)}\\
&\le aL \cdot w(f),
\end{aligned}$$

which proves the claim. □

The above estimate allows us to show that – in the Lipschitz case – the least fixed point of the domain theoretic Picard operator has width 0, i.e. solves the initial value problem, as shown in Lemma 3.7.

Proposition 4.3. *Let $f_{n+1} = P_v(f_n)$ for $n \in \mathbb{N}$. Then $w(f_n) \le c^n w(f_0)$. In particular, $f = \bigsqcup_{n\in\mathbb{N}} f_n$ satisfies $P_v(f) = f$ and $w(f) = 0$.*

Proof. Follows immediately from $aL < c < 1$ by induction. □

In order to actually be able to compute the integrals, we now add approximations to v to the picture, the basic idea being that every continuous vector field can be approximated by a sequence of step functions (i.e. functions taking only finitely many values), which allows us to compute the integrals involved in calculating the approximations to the solution effectively. The key property which enables us to use approximations also to the vector field is the continuity of the mapping $v \mapsto P_v$.

Lemma 4.4. *$P : V \to V, u \mapsto P_u$ is continuous.*

Proof. Suppose $u = \bigsqcup_{n\in\mathbb{N}} u_n$ and $f \in D$. Then $(\lambda x.u_n^-(x, f(x)))_{n\in\mathbb{N}}$ (resp. $(\lambda x.u_n^+(x, f(x)))_{n\in\mathbb{N}}$ is a decreasing (resp. increasing) sequence of functions which converge to $\lambda x.u^-(x, f(x))$ (resp. $\lambda x.u^+(x, f(x))$ pointwise. The claim now follows from the monotone convergence theorem. □

This continuity property allows us to compute solutions to the classical initial value problem by means of a converging sequence of approximations of v.

Theorem 4.5. *Suppose $v = \bigsqcup_{n\in\mathbb{N}} v_n$ and $f_{n+1} = P_{v_n}(f_n)$ for $n \in \mathbb{N}$. Then $f = \bigsqcup_{n\in\mathbb{N}} f_n$ satisfies $f = P_v(f)$ and $w(f) = 0$.*

Proof. Follows from Theorem 3.6 and continuity of $u \mapsto P_u$ by the interchange-of-suprema law (see e.g. [4, Proposition 2.1.12]). □

We have seen that the Lipschitz condition on the vector field v ensures that the approximations of the solution converge exponentially fast (Proposition 4.3). If we now approximate the vector field, the speed of convergence of the approximations to the solution will depend on the speed of convergence of the approximations of the vector field.

We use the following result to estimate the speed of convergence when the vector field is approximated:

Lemma 4.6. *Suppose $u \in V$ and $f \in D$. Then $w(P_u(f)) \leq aL \cdot w(f) + a \cdot w_v(u)$.*

Proof. This is just a matter of calculating

$$
\begin{aligned}
w(P_u(f)) &= \sup_{t\in[-a,a]} \int_0^t u^+(x, f(x)) - u^-(x, f(x))dx \\
&\leq \sup_{t\in[-a,a]} \int_0^t v^+(x, f(x)) - v^-(x, f(x)) + w_v(u)dx && \text{(Def'n of } w_v(u)) \\
&\leq aL \cdot w(f) + aw_v(u) && \text{(Lemma 4.2)}
\end{aligned}
$$

□

As a corollary we deduce that the approximations converge exponentially fast, if the approximations of the vector field do so, too.

Proposition 4.7. *Suppose $v = \bigsqcup_{n\in\mathbb{N}} v_n$ and $f_{n+1} = P_{v_n}(f_n)$. Then $w(f_n) \leq c^n \cdot w(f_0)$ provided $w_v(v_n) \leq c^n \cdot 2M(c - aL)$.*

Proof. We proceed by induction on n, where there is nothing to show for $n = 0$. Given the estimate for $n \geq 0$, we obtain

$$\begin{aligned}
w(v_{n+1}) &= w(P_{v_n}(f_n)) \\
&\leq aL \cdot w(f_n) + a \cdot w_v(v_n) \\
&\leq c^n \cdot aL \cdot w(f_0) + c^n \cdot 2M(c - aL) && \text{(Ind'n Hypothesis)} \\
&= c^n \cdot 2aLK + c^n \cdot 2aM(c - aL) && (f_0 = \lambda t.[-K, K]) \\
&= c^n \cdot 2aLK + c^{n+1} \cdot 2aM - c^n \cdot 2a^2 LM \\
&= c^{n+1} \cdot 2K - c^{n+1} \cdot 2(K - aM) + c^n \cdot 2aL(K - aM) \\
&\leq c^{n+1} \cdot 2K - c^{n+1} \cdot 2(K - aM) + c^{n+1} \cdot 2(K - aM) \\
&\leq c^{n+1} \cdot 2K = c^{n+1} \cdot w(f_0),
\end{aligned}$$

as required. □

Given a representation of v in terms of step functions, Theorem 4.5 gives rise to an algorithm for computing the solution of the initial value problem. Our next goal is to give an estimate of the algebraic complexity of the algorithm.

5 Algebraic Complexity

In order to give an estimate for the algebraic complexity of the algorithm induced by Theorem 4.5, we need to consider the representations of the functions involved in calculating the approximations. Here, we consider approximations by piecewise constant and piecewise linear functions as in [12, 11].

Definition 5.1 (Step Functions). Suppose $A \subseteq [-a, a]$, $B \in \mathbb{IR}$, $C \in \mathrm{I}[-K, K]$ and $f^-, f^+ : A \to \mathbb{R}$ are linear. We consider the following types of step functions, where S^o is the interior of a set S:

1. *Linear single step functions* of type $[-a, a] \to \mathbb{IR}$:

$$A \searrow [f^-, f^+] : [-a, a] \to \mathbb{IR}, \quad x \mapsto \begin{cases} [f^-(x), f^+(x)] & x \in A^o \\ \mathbb{R} & \text{o/w} \end{cases}$$

2. *Simple single step functions* of type $[-a, a] \to \mathbb{IR}$:

$$A \searrow B : [-a, a] \to \mathbb{IR}, \quad x \mapsto \begin{cases} B & x \in A^o \\ \mathbb{R} & \text{o/w} \end{cases}$$

3. *Simple single step functions* of type $[-a, a] \times \mathrm{I}[-K, K] \to \mathbb{R}$:

$$A \times B \searrow C : [-a, a] \times \mathrm{I}[-K, K] \to \mathbb{R}, \quad (x, Y) \mapsto \begin{cases} C & x \in A^o \text{ and } Y \ll B \\ \mathbb{R} & \text{o/w} \end{cases}$$

4. A *simple* (resp. *linear*) *step function* is a finite join of simple (resp. linear) single step functions. The number of linear (resp. simple) single step functions in a linear (resp. simple) step function is denoted by $\mathcal{N}(\,\cdot\,)$.

Note that strictly speaking we should consider representations of step functions when considering $\mathcal{N}(f)$. Since we never consider two or more different representations of the same function, we do not make this distinction for ease of presentation.

If the vector field u is a sup of simple step functions and f is a linear step function, the function $\lambda x.u(x, f(x))$ is simple, hence its integral will be linear again, as in [11, Corollary 4.3].

Proposition 5.2. *Suppose u is a simple step function and f is a linear step function. Then $P_u(f)$ is a linear step function, and can be computed in $\mathcal{O}(\mathcal{N}(u) \cdot \mathcal{N}(f))$ steps.*

Proof. Clearly $s = \lambda x.u(x, f(x))$ is a simple step function, if u is simple. Computing s we need to match every simple single step function in u against every linear single step function in f, which can be done in $\mathcal{O}(\mathcal{N}(u) \cdot \mathcal{N}(f))$ steps. Taking integrals, we obtain a linear step function $P_u(f) = \lambda t. \int_0^t s(x)dx$, which can be computed in $\mathcal{O}(\mathcal{N}(s))$ steps, hence the complexity bound on $P_u(f)$. ☐

Note that the previous lemma shows that it suffices to consider data types for linear and simple step functions in order to formulate the algorithm of Theorem 4.5. However, the number of single step functions needed to represent $P_u(f)$ is also quadratic in general:

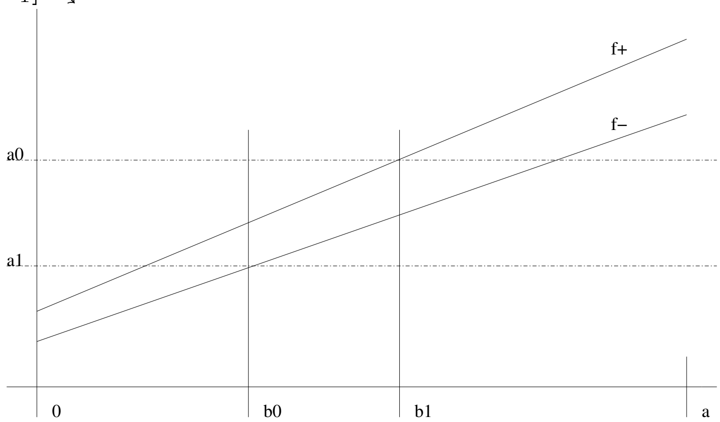
Remark 5.3. If u is a linear step function and f is a simple step function, then $\mathcal{N}(P_u(f)) \in \mathcal{O}(\mathcal{N}(u) \cdot \mathcal{N}(f))$. In particular, if $v = \bigsqcup_{n \in \mathbb{N}} v_n$ and $f_{n+1} = P_{v_n}(f_n)$, then

$$\mathcal{N}(f_{n+1}) \in \mathcal{O}(\mathcal{N}(v_0) \dots \mathcal{N}(v_n)),$$

provided each v_n is a simple step function.

The blow-up of the number of single step functions needed to represent f_n is due to fact that the partition of $[-a, a]$, which is induced by $P_u(f)$ is in general finer than the partition induced by u.

The refinement of partitions happens when computing $u(x, f(x))$, since the linearity of f cuts the partition induced by u, as illustrated for $w = [-a, a] \times [a_0, a_1] \searrow C$ and $f = [-a, a] \searrow [f^-, f^+]$; we obtain $\lambda t.w(t, f(t)) = [b_0, b_1] \searrow C$ where $C \in \mathbb{IR}$ is some interval.



The blow up in the number of single step functions can be avoided if we work with simple step functions only. The key idea is to transform the linear step function $P_u(f)$ into a simple step function before computing the next iterate: on every interval, replace the upper (linear) function by its maximum and the lower function by its minimum. We now develop the technical apparatus which is needed to show that the approximations so obtained still converge to the solution. This is achieved by making the partitions of the interval $[-a, a]$ induced by step functions approximating the solution and those approximating the vector field explicit.

Definition 5.4 (Partitions). 1. A *partition* of $[-a, a]$ is a finite sequence $(q_0, \dots, q_n)$ of real numbers such that $-a = q_0 < \cdots < q_n = a$; the set of partitions of $[-a, a]$ is denoted by $\mathcal{P}$.

2. The *norm* $|Q|$ of a partition $Q = (q_0, \dots, q_n)$ is given by $|Q| = \max_{1 \le i \le n} q_i - q_{i-1}$.

3. A partition $Q = (q_0, \dots, q_n)$ *refines* a partition $R = (r_0, \dots, r_k)$ if $\{r_0, \dots, r_k\} \subseteq \{q_0, \dots, q_n\}$; this is denoted by $R \le Q$.

4. If $f = \bigsqcup_{i \le n} [a_i^-, a_i^+] \searrow [f_i^-, f_i^+]$ (resp. $u = \bigsqcup_{i \le n} [a_i^-, a_i^+] \times [b_i^-, b_i^+] \searrow [c_i^-, c_i^+]$) is a step function, the *partition induced by* f (resp. w) is the

unique partition $(q_0, \ldots, q_k)$ with $\{q_0, \ldots, q_k\} = \{a_i^-, a_i^+ \mid i \leq k\}$; this partition is denoted by $Q(f)$ (resp. $Q(u)$).

Note that the partition induced by a step function depends on the representation of the step function; since we never consider different representations of the same function, we allow ourselves to blur this distinction.

To a partition of $[-a, a]$ we now associate a functional, which transforms continuous functions into simple step functions.

Definition 5.5. Suppose $Q = (q_0, \ldots, q_n) \in \mathcal{P}$.

1. If $g : [-a, a] \to \mathbb{IR}$ is continuous, the *envelope* [10] of g is defined by
$$\mathrm{env}(g) = [\liminf(g^-), \limsup(g^+)]$$
2. The *flattening functional* $F_Q : ([-a, a] \to \mathbb{IR}) \to ([-a, a] \to \mathbb{IR})$ associated with Q is defined by
$$F_Q(f) = \mathrm{env}(\bigsqcup_{1 \leq i \leq n} [q_{i-1}, q_i] \searrow [\inf_{x \in [q_{i-1}, q_i]} f^-(x), \sup_{x \in [q_{i-1}, q_i]} f^+(x)])$$
for $f = [f^-, f^+] : [-a, a] \to \mathbb{IR}$.

Note that a simple step functions is in general undefined at the partition points. Taking the envelope just ensures that the resulting function is defined on the closure of the domain of definition of the step function, i.e. at all partition points, without affecting continuity.

Lemma 5.6. *$F_Q(f)$ is continuous, if f is continuous.*

Proof. Immediate from the fact that step functions are continuous. □

When constructing a solution of a given initial value problem, we want to apply the flattening functional at every stage of the construction to transform a linear step function (obtained by integration) into a simple step function. We still need to show that with the flattening operation the sequence still converges to the solution of the IVP; the following lemma helps us to establish this fact.

Lemma 5.7. *Suppose $(Q_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{P}$ with $\lim_{n \to \infty} |Q_n| = 0$. Then $\bigsqcup_{n \in \mathbb{N}} F_{Q_n} = \mathrm{id}$.*

Proof. Let $x \in [-a, a]$. We construct a sequence $([q_n^-, q_n^+])_{n \in \mathbb{N}}$ of intervals such that

1. $x \in [q_n^-, q_n^+]$ for all $n \in \mathbb{N}$

2. q_n^-, q_n^+ are consecutive partition points of Q_n.

We have $q_n^+ - q_n^- \to 0$ as $n \to \infty$ since $|Q_n| \to 0$ as $n \to \infty$. The claim follows from lower (resp. upper) semi continuity of f^- (resp. f^+). $\square$

The last lemma puts us in the position to show that the application of the flattening functional at every stage of the construction does not affect the convergence of the iterates to the solution.

Theorem 5.8. *Suppose $v = \bigsqcup_{n\in\mathbb{N}} v_n$ where the v_n's are simple step functions with $\lim_{n\to\infty} |Q(v_n)| = 0$. If $f_{n+1} = F_{Q(v_n)}(P_{v_n}(f_n))$, then $f = \bigsqcup_{n\in\mathbb{N}} f_n$ satisfies $f = P_v(f)$.*

Proof. Follows from the interchange-of-suprema law (see e.g. [4, Proposition 2.1.12]), the previous lemma and Theorem 4.5. $\square$

In the following, we investigate the speed of convergence of the f_n's constructed in the previous lemma and investigate the algebraic complexity of computing iterates. It turns out that the speed of convergence is essentially not changed by applying the flattening functional at every step of the construction; this result hinges on the following estimate:

Lemma 5.9. *Suppose $u \in V$ and $Q \in \mathcal{P}$, $f \in D$. Then $w(F_Q(P_u(f))) \leq aL \cdot w(f) + a \cdot w_v(u) + 2\frac{K}{a}|Q|$.*

Proof. In view of Lemma 4.6 and Lemma 3.4, it suffices to show that $w(F_Q(g)) \leq w(g) + 2\frac{K}{a}|Q|$ if $g : [-a, a] \to \mathbb{IR}$ and both g^+, g^- satisfy a Lipschitz condition with Lipschitz constant K/a. Suppose $x \in [-a, a]$ and choose two consecutive partition points q^-, q^+ of Q such that $x \in [q^-, q^+]$. Since upper (resp. lower) semi continuous functions attain their suprema (resp. infima) on compact intervals, there are $x^-, x^+ \in [q^-, q^+$ such that, for all $x \in [q^-, q^+]$, we have $(F_Q(g))^-(x) = g^-(x^-)$ and $(F_Q(g))^+(x) = g^+(x^+)$. Thus,

$$
\begin{aligned}
&(F_Q(g))^+(x) - (F_Q(g))^-(x) \\
&= |g^+(x^+) - g^-(x^-)| \\
&= |g^+(x^+) - g^+(x) + g^+(x) - g^-(x) + g^-(x) - g^-(x^-)| \\
&\leq \frac{K}{a}|x^+ - x| + |g^+(x) - g^-(x)| + \frac{K}{a}|x - x^-| \\
&\leq 2\frac{K}{a}|Q| + w(g),
\end{aligned}
$$

as required. $\square$

Using the last lemma, we can now estimate the speed of convergence as follows:

Proposition 5.10. *Suppose $v = \bigsqcup_{n\in\mathbb{N}} v_n$ with $w_v(v_n) \leq c^n \cdot M(c - aL)$ and $|Q(v_n)| \leq c^n \cdot \frac{a}{2}(c - aL)$. Then $w(f_n) \leq c^n w(f_0)$ if $f_{n+1} = F_{Q(v_n)}(P_{v_n}(f_n))$ for all $n \geq 0$.*

Proof. We use induction. The estimate is evident for $n = 0$; for the inductive case one calculates

$$\begin{aligned}
w(f_{n+1}) &\leq aL \cdot w(f_n) + a \cdot w_v(v_n) + 2\frac{K}{a} \cdot |Q(v_n)| \\
&\leq c^n \cdot 2aLK + c^n \cdot aM(c - aL) + c^n \cdot K(c - aL) && \text{(Ind'n Hypothesis)} \\
&= c^n \cdot 2aLK + c^{n+1} \cdot aM - c^n \cdot a^2LM + c^{n+1} \cdot K - c^n \cdot aLK \\
&= c^{n+1} \cdot 2K - c^{n+1} \cdot (K - aM) + c^n \cdot aL(K - aM) \\
&\leq c^{n+1} \cdot 2K - c^{n+1} \cdot (K - aM) + c^{n+1}(K - aM) && (aL < c) \\
&= c^{n+1} \cdot 2K = c^{n+1} \cdot w(f_0),
\end{aligned}$$

as required. □

The last lemma shows, that flattening does not affect the modulus of convergence. We now return to question of algebraic complexity, and show, that computing iterates is still quadratic, but the number of step functions needed to represent the iterates only grows with the number of step functions needed to represent the approximations of the vector field.

We begin with noting that iterates can still be computed by a quadratic algorithm. Before stating the result, we need to extend the notation $\mathcal{N}(\,\cdot\,)$ to functions, which are envelopes of step functions.

Notation 5.11. If $f = \text{env}(g)$ is the envelope of a simple step function g, we put $\mathcal{N}(f) = \mathcal{N}(g)$.

That is to say, the envelope of a join of n step functions is represented by n step functions. We can now deduce:

Proposition 5.12. *Suppose u is a simple step function and f is the envelope of a step function. Then $F_Q(u)(P_u(f))$ is the envelope of a simple step function, and can be computed in $\mathcal{O}(\mathcal{N}(f) \cdot \mathcal{N}(u))$ steps.*

Proof. Follows from Proposition 5.2, since flattening can be done in $\mathcal{O}(\mathcal{N}(u))$ steps. □

In order to give an estimate on the number of step functions required to represent the approximations, we need the following lemma:

Lemma 5.13. *Suppose $Q(f) \leq Q(u)$. Then $Q(P_u(f)) = Q(u)$.*

Proof. Because the partition of $\lambda x.P_u(x, f(x))$ is the partition of u in case $Q(u)$ refines $Q(f)$. $\square$

We conclude with an estimate on the number of step functions needed to represent the iterates.

Proposition 5.14. *Suppose $v = \bigsqcup_{n\in\mathbb{N}} v_n$ where $Q(v_n) \leq Q(v_{n+1})$. Then $\mathcal{N}(f_{n+1}) = \mathcal{N}(v_n)$ for all $n \in \mathbb{N}$, if $f_{n+1} = F_{Q(v_n)}(P_{v_n}(f_n))$ for all $n \in \mathbb{N}$.*

Proof. Follows immediately from the previous lemma using induction; note that $Q(v_0)$ necessarily refines $Q(f_0)$ since $Q(f_0) = (-a, a)$. $\square$

6 Experimental Results

This section briefly reports some experimental results obtained by implementing the algorithm underlying Theorem 5.8. In the examples, we have chosen $a = 1/2$; the n-th iterate of the vector field subdivides this interval into 2^n partitions of the same length. It is immediate that the approximations can be represented using intervals with rational endpoints, if this is true of the vector field. The task of computing with rationals was delegated to the GMP library [2]. Using our first prototype, we conducted experiments with the equations $f'(x) = 2f(x) + 1$, $f'(x) = 2x$ and $f'(x) = 4xy + 1$, all three equations with the initial condition $f(0) = 0$. It is routine to check that Assumption 4.1 is satisfied in this setup. The experimental results are summarised in the tables below; n denotes the number of iterates we have computed.

$f'(x) = 4xf(x) + 1, f(0) = 0$

n	Accuracy	Time
10	0.0134	0m0.518 s
15	0.00043	0m2.573s
20	1.35756e-05	2m22.067s

$f'(x) = 2f(x) + 1, f(0) = 0$

n	Accuracy	Time
10	0.00882	0m0.073s
15	0.000281	0m2.157s
20	8.79667e-06	2m28.899s

$f'(x) = 2x, f(0) = 0$

n	Accuracy	Time
10	0.000976086	0m0.038s
15	3.05171e-05	0m0.990s
20	9.53674e-07	1m54.007s

The algorithm was implemented in entirely unoptimised form and tested on a 700 MHz Pentium 3 with 384 MB of RAM; the times are wall clock times. The large increase in time when computing 20 iterates is partially due to paging.

7 Acknowldegements

This work has been supported by EPSRC in the UK and the EU project "APPSEM-II".

References

[1] AWA. A software package for validated solution of orinary differential equations. `www.lsi.upc.es/~robert/mirror/interval-comp/intsoft.html`.

[2] *The GNU multi precision library.* `www.swox.com/gmp/`.

[3] O. Aberth. Computable analysis and differential equations. In *Intuitionism and Proof Theory, Studies in Logic and the Foundations of Mathematics*, pages 47–52. North-Holland, 1970. Proc. of the Summer Conf. at Buffalo N.Y. 1968.

[4] S. Abramsky and A. Jung. Domain theory. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3. Clarendon Press, 1994.

[5] J. P. Aubin and A. Cellina. *Differential Inclusions.* Spinger, 1984.

[6] V. Brattka. Computability of banach space principles. Informatik Berichte 286, FernUniversität Hagen, Fachbereich Informatik, June 2001.

[7] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory.* Springer, 1998.

[8] J. P. Cleave. The primitive recursive analysis of ordinary differential equations and the complexity of their solutions. *Journal of Computer and Systems Sciences*, 3:447–455, 1969.

[9] E. A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations*. McGraw-Hill, 1955.

[10] A. Edalat. Interval derivative. Presented at MFPS XIX, 2003. `www.doc.ic.ac.uk/~ae/papers/MFPS2.ppt`.

[11] A. Edalat, M. Krznarić, and A. Lieutier. Domain-theoretic solution of differential equations (scalar fields). In *Proceedings of MFPS XIX*, volume 83 of *Electronuc Notes in Theoretical Computer Science*, 2003. Full paper in `www.doc.ic.ac.uk/~ae/papers/scalar.ps`.

[12] A. Edalat and A. Lieutier. Domain theory and differential calculus (Functions of one variable). In *Seventh Annual IEEE Symposium on Logic in Computer Science*. IEEE Computer Society Press, 2002. Full paper in `www.doc.ic.ac.uk/~ae/papers/diffcal.ps`.

[13] A. Grzegorczyk. Computable functionals. *Fund. Math.*, 42:168–202, 1955.

[14] A. Iserles. *Numerical Analysis of Differential Equations*. Cambridge Texts in Applied Mathematics. CUP, 1996.

[15] Ker-I Ko. On the computational complexity of ordinary differential equations. *Inform. Contr.*, 58:157–194, 1983.

[16] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis*. Dover, 1975.

[17] N. Th. Müller and B. Moiske. Solving initial value problems in polynomial time. In *In Proceedings of the 22th JAIIO - Panel'93*, pages 283–293, Buenos Aires, 1993.

[18] R.E. Moore. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, 1966.

[19] K. R. Jackson N. S. Nedialkov and G. F. Corliss. Validated solutions of initial value problems for ordinary differential equations. *Applied Mathematics and Computation*, 105:21–68, 1999.

[20] M. B. Pour-El and J. I. Richards. *Computability in Analysis and Physics*. Springer-Verlag, 1988.